Top Research Productivity and its Persistence*

A Survival Time Analysis for a Panel of Belgian Scientists

Stijn Kelchtermans♣ and Reinhilde Veugelers♦

Abstract

The paper contributes to the debate on cumulative advantage effects in academic research by examining top performance in research and its persistence over time, using a panel dataset comprising the publications of biomedical and exact scientists at the KU Leuven in the period 1992-2001. We study the selection of researchers into productivity categories and analyze how they switch between these categories over time. About 25% achieves top performance at least once, while 5% is persistently top. Analyzing the hazard to first and subsequent top performance shows strong support for an accumulative process. Rank, gender, hierarchical position and past performance are highly significant explanatory factors.

JEL-classification: J24 ; L31 ; O31 ; O32
Keywords: economics of science; research productivity; hazard models

* This paper benefited from the comments of Paula Stephan during her fellowship at the KU Leuven in the spring of 2005. We also gratefully acknowledge comments from participants at the Conference in Tribute to Jean-Jacques Laffont at IDEI/Toulouse, PAI meeting in Mons, EXtra Workshop in Lausanne, St Anna School in Pisa, and more particularly J. Mairesse, B. van Ark, M. Dewatripont, F. Verboven, G. Friebel B. Sampat, B. Hall, F. Murray, G. Dosi, M. Baguës, B. Bijnens. Financial support of the Belgian Federal Science Policy Office (Interuniversity Attraction Poles P5/26.A)) and the KU Leuven (OT/04/07A) is gratefully acknowledged.
♣ KU Leuven, Naamsestraat 69, 3000 Leuven, Belgium, stijn.kelchtermans@econ.kuleuven.be
♦ KU Leuven, European Commission (BEPA) Brussels, Belgium and CEPR, London, reinhilde.veugelers@econ.kuleuven.be

1 Introduction

What makes someone a top researcher? Why do (some) top performers manage to sustain their high productivity level while others peak in scientific output only sporadically or never? Why does a substantial part of academics hardly ever publish anything? Is this due to intrinsic qualities that are largely fixed before a researcher writes her first paper? Do exogenous factors, like gender, age, scientific discipline or organizational structure explain/contribute to differences in scientific performance and persistence of such differences? Or does the research system endogenously establish persistence of productivity differences by favoring the better researchers?

Insights in the factors that drive differences in research performance and its dynamics have important policy implications. Policy makers are increasingly assessing research performance. The use of publication- and citation counts as instruments for evaluation of individual scientists within research institutes as well as for funding decisions for research labs and universities as a whole is becoming more widespread. Furthermore, the allocation of research funding is increasingly being driven by criteria of scientific excellence, resulting in a concentration of more funds in fewer hands. Yet, there are few academic studies on what drives research productivity. A body of empirical research has recently emerged that attempts to pin down the determinants of scientific productivity, both at the level of the individual researcher and the more aggregate, institutional level. However few studies have addressed the skewed distribution of research productivity, with its concentration in a limited number of top researchers and their persistency over time.

This paper studies top research performance and its persistency over time, using a recent panel data set from the KU Leuven. We first explore the selection of researchers into productivity categories (top, medium, low), using a clustering analysis. To assess persistency, we construct mobility matrices; showing the moves to and from the "top performance" category. Second, we analyze the determinants of top research performance and its dynamics in more detail. We employ a duration model to study the factors that influence the hazard for a researcher to achieve a first and subsequent top performance level, taking into account time-varying covariates and checking for the influence of past (top) performance. We complement the hazard analysis by a logit analysis of the characteristics determining (persistent) top performance.

The data set is a panel comprising ten years of publication data 1992-2001, which allows the separation of age and cohort effects. The data set allows to control for several factors that are expected to play a role in scientific publication activity, such as scientific discipline, gender, age, tenure, rank, hierarchical position (head of team) and seniority. Furthermore, the record of each researcher contains teaching load and other administrative duties.

The remainder of this paper is organized as follows. The next section presents a brief literature survey. Section 3 provides information on the data and how it was assembled from different sources. In section 4 we establish the existence of (persistent) productivity cohorts, while in section 5 we analyze the data in search of the characteristics determining the process towards (persistent) top performance. In section 6 we conclude and touch upon directions for further research.

2 Literature Survey

Most existing studies concentrate on the effects of individual determinants of academic productivity. The aspect of individual productivity that has received most attention is the dynamics of research productivity over the **life cycle of the researcher**. The earlier studies on US data (e.g. Bernier et al. (1975) and Cole (1979)), find a curvilinear relationship between age and both quality and quantity of scientific productivity. A limitation of these earlier studies is their use of cross-sectional data, which does not allow to disentangle age from experience and cohort effects (Stephan 1996). Levin and Stephan (1991), using longitudinal data of American scientists, find that life cycles effects are present in five of the six areas of physics and earth sciences studied, with publishing activity initially increasing and then declining somewhere in mid-career.

Gender differences in scientific productivity are another line of attention of researchers. Several studies have found that female scientists publish at lower rates than male scientists. Using a sample of American biochemists, Long (1993) finds that sex differences in the number of publications and citations are bigger during the first decade of the career but are reversed later. He attributes the lower productivity of females to their overrepresentation among non-publishers and their under representation among the extremely productive.

In a more recent study of French condensed matter physicists, Mairesse and Turner (2002) analyze the impact of **age, gender and education** on research productivity. They confirm a quadratic relation between the age of the scientists and the number of publications. They also find significant positive effects for males and graduates from the French Grande Ecoles. Their results also indicate a positive time trend, suggesting that there has been a wider and faster access to publication.

In view of the significance of team effort in science, it is important to assess **collective effects on individual productivity** as also Stephan (1996) argues. Early research in the USA found researchers at **prestigious departments** to be more productive and cited than their colleagues in lower-ranked universities (Cole and Cole 1973). Also Turner and Mairesse (2002) provide evidence that the quality of other researchers belonging to the laboratory is a crucial variable for explaining individual productivity.

To summarize, existing studies assessing individual research productivity have indicated the importance of individual characteristics like age, cohort and gender as well as collective characteristics of the laboratory or department to which the researcher belongs, whilst controlling for scientific discipline idiosyncrasies.

Most of the studies to date aim at explaining average productivity profiles, ignoring the often skewed distribution of research productivity, with many researchers non-active and a few researchers accounting for the bulk of the publications. Furthermore, they analyze publications in a cross section or a short period of time, not allowing to properly account for time persistence of productivity patterns. It is particularly the latter phenomenon we focus on in this paper, more specifically persistency of top research productivity. This issue of **persistence of research productivity profiles** is underexplored in the literature.

Why would we expect top research productivity to persist or not? First, ***talent*** is important in determining research productivity. Top research may require a "magic gland" (Stephan and Levin 1992), a special edge, an innate ability. Those pre-determined differences are unevenly spread in the population. Those who have it, are always productive, those without, never see their careers take off and flourish. Hence, a differential distribution of talent within the scientific community will lead to research productivity differences which persist over time. Of course, ***luck*** also enters the picture, especially when explaining the occurrence of "hits" following a scientific discovery. Although luck enters in a variety of forms and is often accompanied by serendipity, it nevertheless predicts a more random and non-persistent top research productivity for the individual researcher.

But next to talent and luck, ***effort*** is a particularly important factor explaining scientific output. When researchers decide on the level of effort to impute, they trade costs and benefits. *Costs of effort* will be lower for the more talented researchers, resulting in an interaction between talent and effort driving (persistency of) top performance. Furthermore, in line with research on firm growth (e.g. Jovanovic (1982), Pakes and Ericson (1998)), one can argue that the effect of talent is not a fixed effect over time, particularly when interacted with effort in a learning perspective. Initially researchers may be uncertain about their talent when they enter the field, but gradually discover their capabilities from being active in the field (as in the passive learning models of e.g. Jovanovic (1982)), and/or from making efficiency enhancing investment (as in the active learning models of e.g. Pakes and Ericson(1998)). Both directions predict that younger and "smaller" researchers will have a higher potential for learning and hence variance in performance over time, while older researchers will have a higher persistence in performance as they are more confident about their talents. Once a researcher learns she is good at it, she will be more motivated to put effort into the process.

Several *benefits* may motivate scientists to exert effort, as argued by Stephan & Levin (1992), Dasgupta & David (1994): monetary rewards, recognition and the "puzzle" joy. These motivational forces may explain persistency in research productivity through a process of **accumulative advantage** where motivation to exert effort depends on past performance. The Matthew effect described by Merton (1968) states that the recognition and monetary value awarded to a scientist's accomplishments depends on his status in the scientific community. Highly productive researchers maintain or increase their productivity because they receive recognition and resources, while those scientists who do not, become less productive. Successful scientists get their work more easily published, get easier citations, research funding, quicker tenure and promotion, higher wages, jobs in more prestigious institutions, better infrastructure, better and more PhD and post-doc students… In summary, being more recognized and having access to more and better resources, more productive researchers will devote more effort to scientific output, leading to persistency in research performance. They may even escape the typically observed non-linear age profile, being less likely to relocate efforts at mid-life. Zuckerman & Merton (1971) argue that this tenacity occurs not only because the successful have accumulated advantage and have become addicted to recognition, but also because they see this as a process to validate the judgments of the scientific community that their capacities have been correctly assessed.

There is a lot of anecdotal empirical evidence on accumulative advantage (see e.g. Stephan & Levin (1992)), but little recent systematic empirical analysis. An empirical analysis on productivity dynamics is presented by Allison and Stewart (1974). They use a cross-section of chemists, physicists, and mathematicians in the US, and find that the highly skewed distributions of productivity among researchers can be explained by a process of accumulative advantages. This inequality becomes increasingly unequal as career age increases. In an extension of this study, Allison et al. (1982) examine cohorts of biochemists and chemists, and they confirm that increasing inequality is observed for counts of publications but not for counts of citations to all previous publications.

3 The Data

The data set we use to assess persistency of high research productivity is a unique panel of 1040 scientists within the fields of biomedical and exact sciences, in the period 1992-2001, employed at the KU Leuven, the largest and oldest university in Belgium[1]. KU Leuven has the ambition to establish a top position in poles of excellence, and have a good performance in the other areas[2]. To this end it allocates research funding to research proposals on the basis of (international) peer review of excellence in research. Also, recruitment and promotion criteria carry a strong research quality requirement. In addition, research output (quantity and quality) of all its academic staff is regularly monitored.

Most researchers in the sample were not employed for the full 10 years: some of them retired or left the university before 2001, others joined after 1992. These entries and exits yield an unbalanced panel and allow examining cohort effects. The data set holds on average 781 researchers per year in an unbalanced panel [3].

We pooled different sources of data sets, combining information from the personnel administration of the KU Leuven with bibliometric data. The dataset contains the following information:

[1] Founded in 1425, the KU Leuven is the oldest and largest university in Flanders and Belgium, encompassing all academic disciplines. About 1.400 tenured professors and more than 3.500 researchers are currently employed at KU Leuven, dealing with a student population of more than 30000 students each year. It has the legal status of a private institution, but receives most of its funding from the Belgian Government, both in a direct and in an indirect, competitive way. The basic public funding of the university, that pays for the salaries of the academic personnel, has remained more of less stable in the last decade, which has resulted in a more or less stable total number of professors at KU Leuven. The funding for research on the other hand has increased continuously. Most of this funding is obtained on a competitive basis: about one quarter is private funding from industry, about half comes from project funding from regional/national and EU governments and about one quarter is from the regional government allocated to the KU Leuven based on its share of regional publications, citations and PhDs. Of this research funding, only a very small part (<5%) is spent on financing academic positions (most is spent on pre- en post docs, infrastructure and expenses).

[2] The KU Leuven has as mission statement in the observed period, to be among the top 25 European research universities in a wide number of scientific disciplines. But it aims to be among the top particularly in those disciplines in which it is already strong: biochemistry, biosciences, biomedical and several disciplines in medicine, among which hematology, oncology and cardiology.

[3] Given our focus on persistence, we restricted the dataset to researchers whom we observe for at least 2 periods. For 97 researchers there was only one observation in the dataset; these were dropped.

- Scientific output (per researcher per year[4]) i.e. publications in ISI journals classified by scientific discipline[5,6];
- Individual and career-related variables (per researcher per year) i.e. gender, age, cohort (year of entry at KU Leuven), years of tenure, rank (assistant professor, associate professor, professor and full professor), seniority in rank, full-time versus part time position;
- Organizational membership at group- (exact versus biomedical sciences), faculty- (e.g. medicine) and department (e.g. microbiology and immunology)[7] level;
- Other information relevant for examining scientific performance, viz. teaching load (*actual* load in year-hours), other administrative duties within KU Leuven, head of a research unit.

Table 1 and table 2 present the distribution of scientists over the respective faculties and disciplines within the Biomedical and Exact Science groups of the KU Leuven. Both groups, each comprising three faculties, are comparable in size. Within the group of Exact Sciences, the faculty of Science (195 professors) and the one of Engineering (214) are the largest. In the group of Biomedical Sciences, the faculty of Medicine (462 professors) clearly dominates in terms of size. The vast majority of all researchers in the panel (93%) are professors and hence combine research and teaching activities. Most of them have tenure (75%). Considering both scientific discipline and organizational membership, allows disentangling faculty & department organizational and management effects from scientific discipline specific influences.

The first column in table 2 shows the final distribution of the researchers in the sample over the disciplines. Table 5 reports year averages for the individual and career-related characteristics of the researchers in the whole sample. About 88% of them are male with an average age of 47.8 years. About one third of the researchers are born after 1956. 38% of the researchers entered as a professor after 1992. With respect to tenure, 76% of the researchers in the sample have tenure in every year that we observe them, 15% never has tenure and 9% switches from a temporary contract to a fixed appointment. We distinguish between four main ranks, with rank 1 the entry level ('assistant professor'[8]) and rank 4 the highest possible rank ('full

[4] "Year" always refers to a "database year" i.e. the year in which the publication was taken up in the ISI records.

[5] A key characteristic of the dataset is that it allows controlling for scientific discipline-specific effects: the KU Leuven-Steunpunt O&O Statistics classified every journal covered by the SCI into one or more twelve high-level disciplines, all within the field of exact or biomedical sciences. This allowed us to assign each scientist the disciplines in which she published a paper in that year. About 23% of all researchers could not be assigned to a main discipline because they did not publish in the period 1992-2001. For all others, we determined a 'main discipline' for each researcher, which is defined as the discipline, taken from the twelve high-level disciplines, in which the researcher has the most publications in the period 1992-2001. For 54 researchers there was a tie i.e. they published an equal amount of papers in at least two disciplines. For them we randomly assigned a main discipline from the tie disciplines.

[6] There is also scientific output that does not fall under the scope of the Science Citation Index of the ISI For instance, the ISI database does not include proceedings, which in some disciplines, like engineering, are an important publication outlet.

[7] While officially the "faculty" as organizational unit is mostly responsible for the teaching programs, and the "department" is the organizational unit for research activities, in practice both levels are intertwined, particularly with respect to recruiting and promotion of researchers. The "Faculty" level has a higher hierarchical position, with the "dean" being a member of the "bureau" that decides on recruitment and promotion, on the basis of "advice" from the departments.

[8] In the remainder of the paper, we refer to ranks by their number.

professor')[9]. One quarter of researchers has the most junior rank, whilst about the same quantity has reached the top of the career ladder. More than 80% of the scientists have a full-time position at the university either in one single contract or in a combination of several positions. The average teaching load for a professor amounts to 4.18 year-hours[10] and increases monotonously with rank.
As shown in the first column of table 4, on average a researcher publishes 2.97 articles per year[11]. But this average has a high standard deviation (4.42). The propensity to publish varies greatly among disciplines: the average number of publications per year ranges from 1.72 for mathematics to 6.34 for biosciences[12] (see table 6). All this reflects the importance of scientific discipline specific effects when examining research productivity. In terms of output quality as measured by citations in a 3-year window, the KU Leuven researchers score above the world average in each of the disciplines (except neurosciences & physics). Finally, publication activity has increased over time in all disciplines as shown in figure 1.

4 Descriptive Analysis

4.1 Performance Profiles: Identifying Top Performance

In order to identify top performance we carried out a clustering of the researchers' yearly publication records. We focus our output analysis on the number of publications only, see infra. In particular, we compared each scientist's number of publications in a given year and discipline with the score of his or her colleagues in the same discipline. This allows to correct for the supra documented discipline specific publication patterns as well as time trends. Every researcher is judged not only by his or her main discipline but, given the degree of multidisciplinarity, for each of the twelve disciplines.

We implemented the following procedure to assign every researcher to one of the 'productivity categories' (top, medium, low). For each year, the scientist's performance within each of the twelve disciplines is compared with colleagues who are active in that discipline, using k-means clustering. A researcher is thus classified into top, medium or low, for each discipline reflecting the productivity of the researcher in the discipline in that year[13].

[9] KU Leuven, like most Belgian universities, offers tenure to assistant professors who successfully pass the judgment of their work in the years following their hiring. After this initial tenure decision, for which young professors are primarily evaluated on their research output as opposed to other activities, they can be promoted through full professor based on their research and teaching performance, as well as duties performed within the university, with the latter typically gaining importance as one progresses through the ranks. Hence, rank 1 is non-tenure; rank 2-4 is tenure.

[10] A year-hour gives the average weekly teaching load in an academic year. One year-hour corresponds to 30 teaching hours.

[11] A publication gets on average 4.63 citations and has an impact measure of 3.04. A scientist collaborates on a paper on average with 4.68 co-authors.

[12] This high degree of heterogeneity across scientific disciplines also holds for other output measures such as citations, impact factors and co-authorship.

[13] In total 12x3x10 clusters are created (12 disciplines, 10 years). Comparing the mean number of publications for the clusters within a discipline using a ranksum test, in 12 cases the null hypothesis of identical means cannot be rejected.

Subsequently, we aggregate across disciplines. The researcher is considered as top in a particular year if she belongs to the 'top cluster' in at least one of the twelve disciplines[14]. If the scientist is not in the top cluster for any of the disciplines, but in the 'medium cluster' for one or more disciplines, then she is considered medium. She is classified as having a low output if she scores low for each discipline. Because the publication process not necessarily adheres to a year-by-year logic, the yearly performance measures are used to construct a performance indicator based on a two-year moving window[15]. This avoids that we label very productive researchers who happen to face a year with many projects in the pipeline but with relatively few actual publications as non-persistent in top-level output. The mechanism for judging researchers' output is illustrated in figure 2. On average 14% of observations are classified as a top performance whereas 28% vs 58% end up in the medium respectively the low productivity categories. This top 14% of observations accounts for 42.08% of all publications in the sample, while the 58% observations representing low performance jointly supply 7.76% of all output, confirming the skewedness in distribution of publications.

4.2 Establishing persistence: Once Top, Always Top?

In order to establish whether persistence effects exist, we look at the mobility of researchers between the productivity clusters. Scientists who 'get trapped' in the low research productivity category may spend their time on other things such as teaching or internal management duties, may not have the intrinsic ability for doing research, or may face a system geared at stimulating the past performers. The latter interpretation attributes persistent low scientific output to a cumulative advantage mechanism that disadvantages researchers with a low initial output. The same cumulative advantage will generate persistency at the top.

To capture researchers' mobility across productivity clusters, we constructed a series of transition matrices. Table 3 shows the mobility matrix for the transition of 1992/1993 to 1993/1994, relating the three productivity clusters (low/medium/top) of 1992/1993 to those of the subsequent 2-year interval. 65.9% of researchers in 1993/1994 belong to the low productivity category. If previous output was irrelevant for explaining current performance, we would see the same distribution of researchers within a subgroup as in the whole sample. Taking the 410 scientists who performed low in 1992/1993, we would expect 270 scientists (65.9%) to display a low output again in 1993/1994. The number of researchers (379) who end up in the low productivity group in 1993/1994 is clearly higher, indicating that past low performance serves as a predictor for future low performance. The same effect occurs at the other end of the productivity spectrum: 73.0% of the high performers (89 scientists) in 1992/1993 repeat their high output in the next year, a higher than expected percentage. And also within the middle performance category, there is a higher than expected probability that researchers remain in this category over the years. The 'low' category is the most immobile category with a lower than expected probability to

[14] So a researcher has in principle twelve opportunities to get into the top cluster for a particular year.

[15] This entails the loss of one observation per researcher. The reported results are not sensitive to a yearly versus two-year moving window.

switch to the middle and a fortiori the high category. The Pearson Chi²-statistic (613.65) confirms that performance levels across these two years are not independent. Similar results are obtained for transition matrices covering the other two-year intervals. We may conclude that researchers tend to be rather immobile in their scientific output, especially the low performers. All this evidence goes against a 'luck' theory for explaining top performance, favoring more the intrinsic qualities, the gradual accumulation and Matthew effect in explaining top performance.

To further analyze this persistence of performance, we measure the persistence of each researcher by counting the number of years in a cluster relative to the years of her employment. In particular, we classify a researcher as *persistent top* if she belongs to the top performance category in every two-year window during the period in which she was employed by the university. Analogously, the 'non-persistent top'-group contains those who belong to the top category at least once but not in every two-year window of their employment. The scientists that do publish but never make it to the top group are labeled 'average' whilst the remaining researchers are 'inactive'. This yields an exhaustive classification of researchers into 4 'persistence categories', as shown in table 4.

Only 49 researchers (5% of the sample[16]) are part of the top productivity category for every two-year window of their employment in 1992-2001. These are the persistent top researchers. These 5% persistent top researchers account for 20.37% of total publications in the sample, confirming again the very skew nature of the productivity distribution. About 20% of the sample belongs to the top category at least once but not persistently. Slightly more than half of all researchers (52%) may be classified as 'average' whilst about 24% of the sample has a blank publication record throughout the observed periods.

Pairwise comparisons using the Wilcoxon ranksum test show that the average number of publications differs significantly for the persistence categories, indicating that we can meaningfully distinguish between them. But the different categories do not only differ significantly in terms of number of publications, but also with respect to the quality of publications. Although we focus our output dimension for identifying top performers on quantity of publications only, we find, in line with other studies, that quantity and quality of publications are correlated. All research output measures for quality (citations, impact measures) decline monotonously when moving from the persistent top researchers to the inactive cohort. Nevertheless, the decline is less outspoken with respect to quality measures than with respect to number of publications.

Splitting the individual and career-related variables by persistency category (see table 5), yields some initial hints with respect to different researcher profiles per persistency category. For example, we see that the representation of the two sexes in the productivity categories is fairly even, although women are slightly more represented in the persistent non-active and in the average category, as in Turner & Mairesse (2003). Women are underrepresented in the top performance category, especially in the non-persistent top performers.

Furthermore, we find little or no support for a curvilinear relationship between age and output on average. We do not find the older age cohorts to be underrepresented among the most productive researchers. This confirms that the

[16] If we had used a one-year window to assess persistence, we would have ended with only 10 individuals, or 1% of the sample, in the persistent top category.

most productive researchers can beat the curvilinear age profile. This is not the case for the unproductive researchers who are overrepresented in the age categories above 50 and especially above 60, pointing at segregation on the basis of age at the bottom of the productivity distribution.

Related to the age results are the ones with respect to rank. There is strong overrepresentation of full professors among the persistent top, while the inactive and average researchers are underrepresented in the full professor rank, but overrepresented in the lowest rank. In addition, as shown by average rank seniority, top researchers tend to spend less time through the three ranks preceding full professor than average and inactive researchers, while inactive researchers tend to stay longer in the less-than-full-professor rank. All this suggests that the promotion procedures in place at the KU Leuven tend to select the more prolific faculty.

Also, inactive and average productive scientists are more likely to have entered professorship recently i.e. after 1992. Recent hires might suffer a disadvantage since they are less likely to have made it already to the top cluster given their shorter employment history in a cumulative process.

The table reveals that it is important to correct for the type of contract: only half of the researchers who don't publish have a full-time position at the university.

As far as research-teaching trade-off is concerned, persistent top researchers tend to have a similar teaching load than the average researchers. Striking is the low teaching load for inactive scientists, but this can at least partially be explained by the higher share of part timers in this category (50% versus 20%) given that they typically have a lower teaching load than people in a fulltime position (1.6 versus 4.8 year-hours). When splitting teaching load by rank, persistent top researchers in all but the second rank do teach less than their colleagues from other productivity cohorts who tend to stick closer to the average teaching load. The most marked difference occurs in rank 4 where persistent top researchers have a substantially lower teaching load, particularly as compared to the average. Also note that junior ranked faculty have a lower teaching load, which should allow them to concentrate more on research activities in the early stage of their career.

5 Multivariate Analysis of Persistence in Top Research Performance

In this section, we take a multivariate look at the process to top performance and its persistence. We use various approaches. In section 5.1 we use a duration analysis to study the determinants of high research productivity and its persistency. Section 5.2 uses a logistic model to explain long-term persistence in research productivity.

5.1. The 'risk' to be top

In this section, we use a duration model approach to determine which independent variables are significantly correlated with the survival time in the non-top research productivity category. This allows discussing which characteristics influence the probability of becoming a top researcher at some point in time. We use the most general proportional hazards model because it is not based on any assumption concerning the shape of the underlying survival distribution as opposed to the range of parametric survival models (exponential, Weibull, loglogistic, lognormal…).

5.1.1 The Cox proportional hazards model

We use the Cox proportional hazards model (1972) where the event of interest is top research performance in a given time period[17]. The model specifies the hazard to be top for the *j*-th individual as the product of a baseline hazard $h_0(t)$, i.e. the hazard when all covariates are equal to zero, and the exponential function of the parameters β_x and regressors x_j:

$$h(t \mid x_j) = h_0(t)\exp\left(x_j \beta_x\right)$$

The baseline hazard is left unspecified, meaning that the model makes no assumptions about the shape of the hazard over time. We opt for this approach since theory does not provide us with a reasonable assumption about the shape of the hazard. The cost of this semi-parametric approach is a loss in efficiency.

Whatever the shape of the hazard function, it is assumed to be the same for everyone. This means that the hazard functions of any two individuals who differ in their characteristics are multiplicative replicas from each other.

We use two types of hazard models. The first one analyses the hazard to ***first top*** performance. In this case, once a top performance is reached, the researcher is removed out of the sample. Hence, for this analysis, the sample consists of all researchers before their first top performance. This allows concentrating on the process towards first top performance, but ignores subsequent observations on performance. Nevertheless, given the skewed distribution of top performance, the overwhelming majority of observations are maintained in this analysis.

Using all information in the data, we also estimate a ***repeated events model***, where we model not only the hazard to first top performance, but to all top performances. We impose a sequential ordering of events: a researcher can only be 'at risk' for her *k*-th top performance if she achieved *k-1* top performances in the past. A key element in our approach is that we allow the risk process underlying top performance to change with the occurrence of previous top performances. The main model estimates a common baseline hazard across event ranks as in the Andersen-Gill counting process[18] model (1982) but with the inclusion of a previous events counter which allows the hazard to differ proportionally between top performances. As a robustness check we also estimate the model with a restricted or 'stratified' risk set i.e. with different baseline hazards per top performance, commonly known as the Prentice-Williams-Peterson model (1981). [19]

---

[17] More specifically, we use the two-year intervals discussed in section 4.1.

[18] The counting process definition of the duration variable uses the time of the (k-1)th top performance as the starting time for each risk interval. This set-up is preferable when the substantive interest is in the evolution of the risk to be top as a function of time since the onset of risk.

[19] This latter model may be estimated using either a counting process or gap time formulation of time, allowing us to check which view of the risk process fits the data best. The gap time approach 'resets the clock'

Given that we work with multiple failure-time data, the failure times pertaining to a single researcher will be correlated, violating the independence of failure times assumption required in traditional survival analysis. This problem is avoided in the first top performance analysis, but taking into account repeated events introduces dependence between failure times pertaining to a single researcher. We account for this by adjusting the covariance matrix of the estimators. In general however, this approach does not fully address the problem since an individual's top performances may be correlated due to a characteristic not being measured, such as ability, instead of being brought about by an accumulative advantage process. Therefore, as an additional robustness check we estimate a model where a latent random effect, or 'frailty', enters multiplicatively on the hazard function: $h_{ij}(t) = h_0(t)\alpha_i \exp(x_{ij}\beta)$. The frailties α_i are unobservable and are assumed to follow a gamma distribution with mean one and variance to be estimated from the data. If the variance differs significantly from zero, then the null hypothesis of no unobserved heterogeneity cannot be maintained.

5.1.2 Censoring and truncation

We define the starting time for the duration variable as the moment when the individual enters as assistant professor at KU Leuven. This marks a natural starting point for the onset of risk and it excludes earlier periods of time when the hazard is zero.

Our observation window ranges from 1992 to 2001. This implies that the majority of researchers in the sample (62%) are 'at risk' prior to 1992 and that their performance data are left censored to the extent that we do not know whether or how many times these scientists achieved top performance i.e. the censoring value is not known. Analogously, 86% of researchers is last observed at censoring time 2001 but continues to be at risk and we are ignorant of their performance beyond this point. While this right censoring is not likely to affect our results, the left censoring is a more important issue.

We will consider all individuals who became a member of the faculty prior to 1992 as cases of 'late entry' in the risk set i.e. we treat these individuals as not observationally at risk of being top, ignoring top performance before 1992. We use the same approach for the right-censored observations, thereby making the assumption that durations are independent of observed entry and exit times. Stated differently, we assume that the censoring is non-informative (Cox and Oakes 1994). We nevertheless include an entry cohort dummy for those individuals having entered before 1992. We will also perform the analysis on the sample excluding the faculty which have entered before 1992.

Finally, the dataset contains observations that are terminated before censoring time. It concerns individuals who left the university or who retired before 2001 (14% of researchers). The precise reason of termination may be important for the same non-informative censoring requirement as above, a condition which is violated if termination is in some way related to survival time. This could be the case if

to zero after each top performance. This definition of time is preferable when the risk process varies as a function of time since the occurrence of the previous top performance.

individuals tend to self-select out of the university once they gain experience about their capacity to deliver as a researcher[20].

5.1.3 Variables

The dependent variable is the hazard to enter the category of top performance, either as a first or repeated event. The categorization of top performance, as defined in section 4.1, is time and discipline specific. This takes care of discipline and time specific effects that may drive the definition of what constitutes a top performance in terms of number of publications required for such top performance.

Which independent variables do we consider to influence this hazard to top performance? The existing studies assessing individual research productivity reviewed in section 2, have indicated the importance of individual characteristics like age, cohort and gender. Also collective characteristics of the laboratory or department to which the researcher belongs have been indicated as important. In line with the existing literature we include the same characteristics, not to check whether they influence research productivity in general, but whether they play a role in explaining the hazard to (first and repeated) top research performance.

Following earlier studies, we include *gender* to check whether females are less likely to deliver (repeated) top performance. Second, we include *age* and *age squared*, with the squared term to check for non-linear age effects, as in previous analysis on research productivity. We expect age to be beneficial for generating top performance, given that it takes time and experience to build an advantage. Furthermore, if there is accumulative advantage, we expect no concavity for age in the repeated top performance analysis, with top performers being able to beat the age decline. To disentangle age from cohort effects, we also include dummies for entry into the sample. The most important *cohort* effect seems to be a marked increase in hiring by the KU Leuven in 1992[21]. There were no other cohort shocks that could be identified[22]. The dummy for entry after 1992 also allows at the same time to correct for left censoring.

All existing studies indicate the importance of controlling for scientific discipline idiosyncrasies. Although our top performance classification is already discipline and time specific, we nevertheless include *scientific discipline* dummies to check whether discipline-specific aspects drive the process towards top performance.

Beyond the traditionally considered variables like age, gender, cohort and scientific discipline, the KU Leuven personnel data set also allows to include a number of other determining variables. A first set of variables reflects the influence of the personnel strategy of the university to reward and provide incentives to its

[20] In support of the argument that the data displays non-informative random censoring, we see that the majority of terminations happens at relatively advanced age (more than 60% of leaving scientists is older than sixty, on a pre-retirement scheme). Furthermore, for the terminations before retirement time it is likely that these people leave because they, for example, got a good outside offer and not because they consider themselves inadequate researchers: there are no significant differences in research performance between the researchers leaving the university before the age of 50 and those of comparable age, rank and discipline, staying. Moreover, the university has not yet a history of firing people for low research performance.

[21] This peak in hiring corresponds mainly to a growing number of retiring faculty that needed to be replaced.

[22] When including a full set of cohort dummies, no significant effects could be detected.

researchers: tenure, rank, and seniority in rank. With respect to *tenure*, we would expect, all else equal, that those researchers up for tenure, will have a higher motivation to provide inputs into the research process. Once tenure is acquired there is less motivation for delivering star performance. The same holds for *rank:* in lower ranks, researchers should have more incentives to put in effort to get promotion. On the other hand, the higher ranks also have more incentives to put in effort to "prove their rank". Since past research performance is taken into account when hiring and promoting, it is likely that top performance will increase the probability of getting a higher rank. To take this endogeneity (at least partly) into account, we include the rank variables with one period lagged relative to top performance. The variable *seniority in rank* should capture increasing pressure to provide effort, the longer a researcher is in his current rank (since the more likely she is to be up for promotion). But again we might expect a non-linearity, once a researcher is far beyond the expected seniority (typically two years), this might reflect a structurally reduced probability to get promotion. Also the more seniority, the higher is the wage and thus the smaller is the marginal benefit from increasing wage with rank. Especially in the end rank (full professor) seniority in rank looses its specific function and will correlate with age.

Beyond the seniority in rank, we also include *seniority as professor*. This variable might be important beyond the seniority in rank, since wages received by professors in Belgium is beyond rank also, and strongly, determined by seniority as professor.

We also include the *organizational unit* to which the researcher belongs at the KU Leuven. This allows capturing the influence of organizational structure and strategy to promote and provide incentives for research, to the extent that these units are responsible for developing a good research environment. It also allows correcting for the impact of spillovers from the quality or prestige of the group to which the researcher belongs. We include "faculty" dummies, since this is the most decisive organization level at KU Leuven in terms of recruitment and promotion decisions. But we also perform analyses with "department" dummies.

Also important is to correct for *fulltime or part time appointment* at the universities, since part time appointments, in our sample mostly occurring at the engineering faculty, are typically for people from industry who are hired and evaluated on teaching rather than research. We include a dummy for fulltime position, but will also report the results for fulltime researchers only.

The inclusion of *actual teaching load* should be able to correct for the lost opportunity time for research when having to teach students. Hence, we expect a negative effect on the hazard for top performance.

Finally, we also have information on whether a researcher heads a research unit. A *head of a unit* has access to resources for research (infrastructure and researchers), which allows him to be more associated with research output in the form of publications in his own name or from his team, as last author. Given that top performers are more likely to become heads of unit, there is an issue of endogeneity. We lag this variable by one period.

5.1.4 Results

a) Hazard to first top performance

We start with discussing the results from the Cox model on first top performance only. The estimates in Table 7 are presented in terms of their effect on the odds to be a top performer. A coefficient smaller (larger) than one, reflects a negative (positive) effect.

Being male increases the odds of displaying first top performance significantly and considerably, by a factor of 2.5. Rank is highly significant and the size of the coefficients suggests that lower ranks have a significantly lower hazard for first top performance as compared to rank 4 (full professor). This may be picking up the accumulative advantage effect that higher ranked professors get more resources, have more incentives to put in effort, and more experience, increasing their probability of realizing a top performance. On top op the rank effect, being a head of unit also considerably and significantly increases the hazard to first top performance. With heads of unit having available the research resources in their unit, they are more likely to be prolific, again supporting the accumulative advantage effect. Once corrected for rank, age has a small and non-significant effect on the hazard and there is no sign of non-linearity in the age effect. Also seniority and tenure have no significant effect[23]. Teaching load is significant and negative, suggesting a substitutive effect between research and teaching, although the magnitude of the effect is small.

The correction for entry cohort after 1992 shows a significantly higher hazard to first top for late entrants, suggesting that the researchers hired before 1992 are not at higher risk for top performance, on the contrary. Also the control variable of having a fulltime position at the university is significant and with a high, positive impact: part timers are indeed less likely to be prolific in research. Both results suggest the importance of correction for entry and fulltime position[24]. Even if we correct for the main discipline of the researcher in the model[25], the membership of the different faculties[26] matters for agriculture, medicine and pharmacy. Looking at the more detailed department membership (estimates not reported) reveals that organizational differences also continue to play a role *within* faculties.

When we split the analysis by group (biomedical, engineering and sciences) the importance of rank holds across groups (results not shown). The engineering faculty shows a positive but slightly concave age effect at the 10% significance level while it remains absent for the other groups. The gender and fulltime effect is driven by the biomedical observations as well as the impact of heading a research unit, while the negative teaching effect occurs mainly in the engineering and science faculties.

To further examine whether the process to first top performance is a gradual process requiring a steady build up of publications or rather a discontinuous burst in

[23] Dropping tenure and seniority in the analysis does not improve the significance of the age coefficients.

[24] The second specification in table 7 reports the results for the sample concentrating on fulltimers only. Most results remain, except for the substitute effect for teaching which looses significance. Also for the subsample with entrants after 1992, the direction and magnitude of most effects are maintained, but the significance for many parameters disappears due to a limited number of observations.

[25] All discipline dummies are relative to the agriculture and environment category. The multidisciplinary category was dropped from the regression (10 researchers).

[26] The faculty of physical education and kinesiology serves as the reference category.

publication output, we report in column (3) the results when including previous research performance. We include a dummy which takes the value of 1 when the researcher belonged in the previous period to the middle performance category, the benchmark being the low performance category. This coefficient turns out to be highly significant and suggests a large effect: researchers in the middle performance category are about 5 times more likely to reach their first top performance in the next period, as compared to researchers in the low performance category, all else equal. This result strongly supports the gradual build up of top performance, as was also found in the mobility matrices reported in section 4. An additional check shows that this effect holds for both male and female researchers. Women are less likely to jump from the low to the top productivity category, but the difference is not significant.

b) Repeated events results

Table 8 presents the same model but this time to predict hazard to all top performances using a repeated events model. These results are particularly interesting to examine the persistency in top performance. When comparing these results with first top performance, we can single out whether the risk process underlying first and subsequent top performance is different[27].

The results in terms of which characteristics are significant are very similar to the results for first top performance, such as the slightly negative impact of teaching load on the hazard to be top. The gender effect remains important and very significant. Also the rank effect remains highly significant with all ranks less likely to perform top although the gap with the full professor rank is smaller in the repeated events regressions. Somewhat surprisingly, the head-of-unit result is less strong in the repeated events analysis. The correction for fulltime position becomes even more important once we take into account subsequent top performances.

To further examine the persistence in research performance, we include in the repeated events analysis, the number of previous top performances (see Table 8, second column). If top performance is a persistent, accumulative process, we would expect this variable to significantly affect the hazard for repeated top performance. Indeed the results indicate a significant and substantial higher hazard for next top performance, the more previous top performances a researcher has acquired. Note that the other covariates are little affected by the inclusion of the variable. In the second specification we test to which extent there is a gradual accumulation of top performance experience, by allowing the effects of first, second, third…top performance on the next top performance to differ. With the exception of the dip in the second top performance, the coefficients indicate an upward trend, suggesting that the more previous top performances the researcher has acquired the higher the hazard for a next top performance, confirming the accumulative nature of the process towards top performance. These findings are confirmed when estimating the stratified Cox model with separate baseline hazards per event rank, as explained in section 5.1.1. With respect to the appropriate view on the risk process, we compared the estimates of the stratified Cox model using both a gap time and a counting process formulation of time. The counting process formulation fits the data best. This implies that the baseline hazard to be top is a function of the total time path since

[27] Given that most of the observations in our sample (83%) pertain to first top performance, it will be difficult to find strong differences.

entry as a professor, rather than being determined by the time since last top performance only.

Table 9 reports the results using a Cox shared frailty[28] model. As discussed in section 5.1.1, if there exists within-individual correlation, the model above is misspecified. The frailty model allows checking whether the accumulative advantage effect coming out of the model is robust when controlling for unobserved heterogeneity. The main parameter of interest is the variance of the frailty terms (θ). The likelihood ratio test for the second model in the table shows that we can reject the hypothesis of no individual heterogeneity. But when correcting for individual fixed effects, all supra reported results remain. More particularly, the effect of previous top performance, although smaller than before (1.17 versus 1.47 as shown in the table), remains sizeable and very significant. The third model in the table reveals that men and women show a different sensitivity to past top performance with respect to the odds of repeating top performance. For a female researcher, each top performance almost triples the odds to be top again (2.51). For the male researchers, these odds are only half (0.43). Note that we control for individual heterogeneity so the different impact of past top performance for men and women cannot be attributed to a systematic difference in, say, unobserved ability. The estimates would suggest that female scientists are more sensitive to the cumulative advantage effect than men. Intuitively, this may be explained as top performance having an even stronger status effect if achieved by researchers operating in an underdog position, which arguably holds for women in science.

Finally, when plotting the baseline hazard (see figure 3), visualizing the shape of the residual risk over time after controlling for all the observable factors, we see that the more time elapses since entry in professorship the less likely it is the researcher will ever reach top performance.

In sum, the estimated models yield interesting findings with respect to the hazard analysis of top performance. In the duration models, the discipline of researchers as well as their mode of employment at the university (fulltime versus part time) turned out to be very important control variables. Further, and in line with previous research, we found a strong gender effect. In contrast, no evidence was found of any age effect nor a significant or sizeable impact of tenure or seniority on top productivity. Nevertheless, career incentives do matter: the likelihood to be a top performing researcher increases with rank. In this respect, heading a research lab also has a positive effect albeit marginally significant in the repeated events regressions. Teaching was found to have a limited but significant substitutive effect on research output. We also characterized the path towards top performance as gradual rather than abrupt. An important finding is the significant and accumulative impact of previous top output on the probability to repeat such an accomplishment, especially for women.

5.2. Probability of being Persistent at the Top

In the previous section we used a hazard analysis to look at the process to first and subsequent top research performance. In this section we use as alternative

[28] The frailty is shared by all observations pertaining to a *single individual* and hence captures within-individual correlation.

approach, a logit analysis of (persistent) top performance to check the robustness of our findings.

The logit analysis examines which characteristics explain the probability to be selected into a productivity cohort, collapsing the time dimension towards top performance. We start in section 5.2.1 by estimating a logit model that explains the membership of the 2-yearly productivity clusters. In section 5.2.2 we focus on the selection into the persistent top category.

5.2.1. Explaining selection into the top research category

We start with an analysis of the probability to be in the high performance category relative to being average or low. The estimated standard errors account for the fact that we have repeated observations for the individuals. The results are shown in table 10. The estimates are reported as risk ratios, relative to the middle & low output category.

The results confirm the hazard analysis of section 5.1.4a. The relative 'risk' of being in the top category versus an average or low output is significantly higher for males than for females. While age and seniority again have no significant impact, higher ranks are significantly more likely to be associated with high research performance, as well as being a head of unit, again supporting the cumulative productivity effect. Also the corrections for discipline, faculty membership and fulltime position, as well the negative effect of teaching load are confirmed.

5.2.2. Explaining selection into the persistent top research category

Next we use the logit approach to analyze the probability of persistent top performance. The logit analysis runs into the problem of a very skewed distribution, with only a limited number of non zero's i.e. persistent top performers (N=49), which makes it hard to find a good predictive model. This is why we also perform the analysis, defining persistency on "good" performance (top and/or average[29]), rather than "top" performance only. Furthermore when we focus on the selection into the persistence categories, we have to average all the variables over the considered time period per researcher. For the time-varying rank dummy variable, we take the highest rank achieved over the considered period, while for the head of unit and tenured position, we evaluate whether the position was ever achieved during the considered period.

As expected, the logistic on the highly skewed persistent top performance gives very bad results, with only the rank effect "surviving" as significant effect. The tenure effect is significantly negative but only applies to researchers with the most junior rank as higher rank scientists are always tenured[30]. The magnitude of the fulltime dummy is explained by a single part time researcher present in the persistent top category. Turning to the less skewed case of persistent good performance, gives more significant results. These results are in line with the repeated hazard analysis of section 5.1.4b. After correcting for scientific discipline and fulltime position, gender and rank are significant characteristics in explaining persistency in good

[29] In addition to being more 'permissive' towards fallbacks in performance, we also allow for a start-up phase: a researcher may initially be part of the low productivity category. Once top performance is achieved she may only fall back to the middle productivity category, not the low one.

[30] If tenure is dropped from the model, the effect is completely absorbed by the first rank.

performance, with males and higher ranked academics having a higher probability of selection into persistent toppers. Also the substitute effect of teaching load again shows up significantly, while heads of unit come out strongly significantly as having a higher probability of selection.

6 Conclusions and Further Research

The paper uses a panel dataset comprising the publications of biomedical and exact scientists at the KU Leuven in the period 1992-2001, to (i) study the selection of researchers into productivity cohorts: and (ii) analyze how they switch between these categories over time, contributing to the debate on cumulative advantage effects in academic research.

We find that productivity cohorts are generally persistent over time. The 'low' category is the most 'trapped' category with a lower than expected probability to switch to the middle and a fortiori the high productivity cohort in later years. About one quarter of the scientists in the sample achieves top performance at least once in the observation period, with five out of a hundred scientists being persistently top.

A hazard model predicting the time towards first top performance confirms the importance of gender, with females being significantly less likely to reach top performance. Age and cohort effects are not significant, but rank and hierarchical position, as well as previous performance are important for explaining the hazard to first top performance, confirming that first top performance is a gradual, accumulative process, as the Matthew effect or an innate ability perspective would predict.

When analyzing subsequent top performances, we again find strong support for the accumulative process, with the hazard to next top performance being significantly and increasingly positively affected by previous top performance. Rank is important not only in predicting first top performance, but also for persistency in top performance, supporting even further the accumulative effect. Also the gender bias remains significant in explaining persistency in top performance, with the dependence on previous top performance in favor of females.

There is some limited but significant and consistent evidence with respect to the substitution effect of teaching load on top research performance. Correction for scientific discipline, full time position and organizational membership is important.

Although the current analysis provides interesting results with respect to top performance and its persistence, it also suggests exciting avenues for further analysis. A first important extension, suggested by the significance of organizational affiliation dummies of the current analysis, is to examine the collective effects in more detail, by specifying characteristics of the research teams and networks to which the researcher is affiliated, and with whom she cooperates in research through co-publications, within KU Leuven but also beyond. The significance of the teaching effect suggests extending the analysis to properly take into account the substitution - or complementarity - among the various output dimensions for researchers: basic research, teaching but also applied research and own commercialization (patents and spin-offs). Also the trade off between quantity and quality of research output can be investigated in more detail. While the current analysis has focused on establishing top performance and its persistence in terms of quantity of publications, we can extend and compare the analysis taking on board the quality of publications dimension, using citation information. Furthermore, by examining whether publications that yield more peer recognition, through citations, are more likely to establish persistence in performance we can study the processes governing the Matthew effect more carefully. And this brings us to perhaps the most important extension suggested by the current analysis, namely to further characterize the accumulative process of persistent top performance. Next to promotion, access to research lab and team infrastructure, research funding is an important component of the accumulative advantage story. For this we need to extend the database with information on research funding received, to examine whether it contributes to establishing and maintaining persistency in research performance. When including funding and other accumulation components in the performance analysis, the endogeneity needs to be properly taken into account, by including an instruments or systems approach.

Finally, an important limitation of this work is the limited scope of the data since we consider the KU Leuven only. Comparison of results with other institutions would be very valuable to improve on institutional heterogeneity.

7 Tables

Table 1: Distribution of Researchers over Organizational Units

Organizational Unit	Freq.	Percent
Group Exact Sciences	486	47.0
Faculty of Science	195	18.8
Faculty of Engineering	214	20.7
Faculty of Applied Bioscience and Engineering	77	7.4
Group Biomedical Sciences	549	53.0
Faculty of Medicine	462	44.6
Faculty of Pharmaceutical Sciences	36	3.5
Faculty of Physical Education & Kinesiology	51	4.9
Total	**1035***	**100.0**

* five people switched between groups and/or faculties in the period 1992-2001 and are not shown in this table.

Table 2: Distribution of Researchers over Disciplines

Main Discipline	Freq.	Percent
None (inactive researchers)	246	23.65
Clinical and Experimental Medicine II (Non-internal Medicine Specialties)	187	17.98
Clinical and Experimental Medicine I (General & Internal Medicine)	152	14.62
Biosciences (General, Cellular & Subcellular Biology; Genetics)	87	8.37
Chemistry	84	8.08
Engineering	75	7.21
Physics	60	5.77
Agriculture & environment	33	3.17
Biology (Organismic & Supraorganismic level)	32	3.08
Biomedical research	27	2.6
Mathematics	24	2.31
Geosciences & space sciences	20	1.92
Neuroscience & behavior	13	1.25
Total	**1,040**	**100.0**

Table 3: Transition matrix 1992/1993 – 1993/1994

Best performance in 1992/1993		Best performance in 1993/1994 low	middle	high	Total
low					
	freq	379	29	2	410
	expected freq	270	88	51.9	410
	%	92.44	7.07	0.49	100
middle					
	freq	25	88	12	125
	expected freq	82.3	26.8	15.8	125
	%	20	70.4	9.6	100
high					
	freq	7	17	65	89
	expected freq	58.6	19.1	11.3	89
	%	7.87	19.1	73.03	100
Total					
	freq	411	134	79	624
	expected freq	411	134	79	624
	%	65.87	21.47	12.66	100
Pearson chi2(4) = 613.6466 Pr = 0.000					

researchers exit (in 92, but not in 93): 26
researchers entry (not in 92, but in 93): 58

Table 4: Research Output by Persistence Category (yearly averages)

Variable	Whole sample	Persist. top	Non-persist. top	Average	Inactive
N	1041	49	206	540	246
%	1.00	0.05	0.20	0.52	0.24
Publications per author	2.97	13.50	6.34	2.07	0.00
	(4.42)				
Co-authors per publication	4.68	4.75	4.54	4.72	n/a
	(3.47)				
Citations per publication	4.63	5.37	4.80	4.50	n/a
	(7.05)				
Impact measure per publication	3.04	3.29	3.11	2.98	n/a
	(2.38)				
(co-)Promoted PhDs	0.25	0.73	0.58	0.18	0.04
	(0.47)				

s.d. between brackets

Table 5: Individual and Career-related Variables by Persistence Category (yearly averages)

Variable	Whole sample	Persist. top	Non-persist. Top	Average	Inactive
N	1041	49	206	540	246
Male	0.88	0.94	0.96	0.86	0.86
	(0.32)				
Age	47.77	47.41	46.56	46.60	51.40
	(8.99)				
% age cohort 1 (age<40)	0.25	0.26	0.27	0.28	0.15
	(0.38)				
% age cohort 2 (40<=age<50)	0.31	0.31	0.34	0.34	0.24
	(0.36)				
% age cohort 3 (50<=age<60)	0.31	0.29	0.31	0.29	0.34
	(0.37)				
% age cohort 4 (60<age)	0.13	0.14	0.08	0.10	0.26
	(0.29)				
Entry cohort*					
% entry cohort 1 (1955-1980)	0.18	0.34	0.20	0.15	0.20
	(0.39)				
% entry cohort 2 (1981-1988)	0.19	0.23	0.30	0.17	0.15
	(0.4)				
% entry cohort 3 (1989-1991)	0.24	0.04	0.25	0.26	0.22
	(0.43)				
% entry cohort 4 (1992-...)	0.38	0.38	0.24	0.42	0.43
	(0.49)				
Years of employment in 1992-2001	7.5	6.63	8.91	7.57	6.34
	(2.88)				
Fulltime at university	0.81	0.97	0.96	0.87	0.50
	(0.38)				
Rank**					
% rank 1 (junior)	0.24	0.10	0.12	0.26	0.36
	(0.03)				
% rank 2	0.22	0.05	0.14	0.25	0.27
	(0.02)				
% rank 3	0.16	0.08	0.18	0.17	0.13
	(0.03)				
% rank 4 (senior)	0.25	0.63	0.39	0.19	0.16
	(0.02)				
Rank seniority	5.75	8.56	6.02	5.24	6.13
	(6.73)				
% rank 1 (junior)	2.41	0.75	1.85	2.27	2.99
	(2.26)				
% rank 2	3.12	0.74	2.08	3.15	3.79
	(2.69)				
% rank 3	4.26	2.04	2.94	4.33	6.46
	(5.01)				
% rank 4 (senior)	12.36	11.87	10.31	12.79	17.11
	(8.53)				
Teaching load (year-hours)	4.18	4.68	4.78	4.12	3.52
	(4.09)				
% rank 1 (junior)	1.48	0.76	1.33	1.43	1.68
	(1.77)				
% rank 2	3.07	3.36	3.22	3.15	2.76
	(3.1)				
% rank 3	4.95	2.91	4.09	5.31	5.37
	(3.67)				
% rank 4 (senior)	7.95	6.06	7.55	8.80	7.93
	(4.38)				

s.d. between brackets

* For 44 researchers this information is missing.

** Only the four main ranks shown. People may be in other ranks which are of lesser concern here (e.g. 'jury member PhD') or may combine one of these other ranks with one of the main ranks.

*** Expected rank seniority for someone in a given rank; not the total number of years scientists tend to spend in each rank. Note that the mean and s.d. of rank seniority by productivity cohort (making abstraction of rank) is conditional on being in one of the four main ranks.

Table 6: Research output by discipline, active researchers only (N=795)

Main discipline	Publications per author	Co-authors per publication	Citations per publication*	(Citations per publication)$_{World}$**	Impact measure per publication
Agriculture & Environment	3.47	3.38	2.17	1.82	1.68
Biosciences (General, Cellular & Subcellular Biology; Genetics)	6.34	5.68	9.44	7.45	6.40
Chemistry	4.85	3.67	3.68	2.77	3.11
Engineering	1.73	3.26	1.62	1.12	1.15
Geosciences & Space Sciences	2.42	3.86	3.42	3.02	2.55
Mathematics	1.72	1.81	1.55	1.00	0.91
Clinical and Experimental Medicine I (General & Internal Medicine)	5.08	5.81	6.46	4.96	3.97
Clinical and Experimental Medicine II (Non-internal Medicine Specialties)	2.91	5.10	3.67	2.83	2.43
Neuroscience & Behavior	3.00	3.22	5.08	5.47	5.32
Physics	5.14	5.11	2.91	3.13	2.91
Biomedical Research	3.65	4.55	6.58	4.01	3.84
Biology (Organismic & Supraorganismic level)	3.58	4.03	5.27	3.25	3.53
Total	**4.01**	**4.74**	**4.83**	**3.40**	**3.36**
(s.d.)	(5.98)	(5.62)	(6.68)		(3.21)

* The average number of citations received by a paper in a 3-year citation window, for papers published by K.U.Leuven researchers in 1992-2001

** The average number of citations received by a paper in a 3-year citation window, for papers published in 1992-2001 at the world level (source: Steunpunt O&O Statistieken, Leuven, using ISI-data)

Table 7: Cox models for first top performance

Coefficient	Cox: first top performance		Cox: first top performance (full-timers only)		Cox:first top performance w/ 1-period lag in performance indicator	
	Haz. Ratio	z	Haz. Ratio	z	Haz. Ratio	z
male	2.45**	(2.37)	2.38**	(2.28)	2.36*	(1.84)
age	0.98	(-0.17)	0.96	(-0.42)	0.95	(-0.35)
age squared	1.00	(-0.57)	1.00	(-0.25)	1.00	(-0.12)
main discipline						
biosciences	2.57**	(2.51)	2.29**	(2.25)	1.10	(0.16)
chemistry	5.85***	(5.19)	4.96***	(4.77)	3.57**	(2.22)
engineering	5.67***	(4.52)	3.98***	(3.39)	5.52***	(3.00)
geosciences	7.56***	(3.77)	6.53***	(3.43)	6.36	(1.61)
mathematics	10.08***	(5.79)	9.08***	(5.50)	13.03***	(4.01)
medicine I	4.29***	(4.52)	3.62***	(4.07)	2.71**	(2.02)
medicine II	4.84***	(5.03)	4.13***	(4.64)	3.91***	(2.93)
neuroscience	9.66***	(4.43)	8.61***	(4.26)	10.20***	(3.25)
physics	7.13***	(4.96)	6.05***	(4.45)	9.62***	(3.81)
biomedical	2.42	(1.55)	2.01	(1.24)	2.19	(1.09)
biology	0.68	(-0.50)	0.64	(-0.57)	0.00***	(-55.72)
faculty membership						
fac of science	1.43	(0.76)	1.48	(0.82)	0.48	(-1.22)
fac of engineering	1.49	(0.86)	1.47	(0.81)	0.96	(-0.07)
fac of agriculture	4.93***	(3.24)	4.55***	(3.09)	2.85*	(1.74)
fac of medicine	2.62**	(2.37)	2.76**	(2.51)	1.70	(1.14)
fac of pharmacy	3.00**	(2.21)	3.18**	(2.34)	1.79	(0.84)
rank						
rank 1 (t-1)	0.15***	(-5.77)	0.16***	(-5.53)	0.15***	(-4.36)
rank 2 (t-1)	0.24***	(-5.86)	0.25***	(-5.82)	0.18***	(-4.35)
rank 3 (t-1)	0.30***	(-4.89)	0.29***	(-4.94)	0.35***	(-3.37)
other rank (t-1)	0.44***	(-2.60)	0.44**	(-2.54)	0.45*	(-1.75)
seniority in rank	1.00	(-0.22)	0.99	(-0.41)	1.02	(0.57)
head of unit (t-1)	1.50**	(2.26)	1.50**	(2.26)	1.63**	(2.06)
fulltime at university	2.47**	(2.16)			1.48	(0.89)
tenured	1.12	(0.26)	0.99	(-0.02)	1.13	(0.30)
entry as professor >= 1992	1.64**	(2.09)	1.58*	(1.86)	1.65	(1.61)
seniority as professor	1.03	(1.19)	1.02	(0.99)	1.01	(0.42)
teaching load	0.96*	(-1.93)	0.96	(-1.54)	0.94*	(-1.86)
medium performance (t-1)					4.85***	(6.49)
N	4616.00		3732.00		4117.00	
ll	-1106.00		-1053.34		-630.15	
chi2	262.99		202.95		9582.55	

base categories: rank4 (rank), faculty of physical educ. & kinesiology (faculty membership), agriculture & environment (main discipline)

* p<0.10, ** p<0.05, *** p<0.01

Table 8: Cox models for repeated top performance

Coefficient	Cox for repeated events		Cox for repeated events, w/ prev. events counter		Cox for repeated events, w/ prev. events counter dummies	
	Haz. Ratio	z	Haz. Ratio	z	Haz. Ratio	z
male	2.53***	(2.80)	1.97**	(2.17)	1.68*	(1.93)
age	0.94	(-0.71)	0.93	(-0.92)	0.95	(-0.96)
age squared	1.00	(-0.04)	1.00	(0.17)	1.00	(0.28)
main discipline						
biosciences	1.35	(0.85)	1.91**	(2.57)	1.42*	(1.67)
chemistry	3.70***	(4.22)	3.00***	(5.08)	1.99***	(4.01)
engineering	5.10***	(4.70)	3.80***	(5.19)	3.01***	(4.90)
geosciences	6.28***	(4.61)	3.26***	(4.24)	2.35***	(3.15)
mathematics	6.66***	(5.21)	4.36***	(5.90)	2.42***	(3.62)
medicine I	2.21***	(2.65)	2.44***	(3.90)	1.83***	(3.24)
medicine II	2.80***	(3.45)	3.07***	(4.95)	2.11***	(4.03)
neuroscience	4.26***	(3.75)	3.47***	(3.48)	2.58***	(4.01)
physics	4.91***	(4.48)	4.10***	(5.13)	3.01***	(4.69)
biomedical	2.37*	(1.88)	2.40**	(2.57)	1.72**	(2.09)
biology	1.34	(0.49)	1.34	(0.66)	1.26	(0.70)
faculty membership						
fac of science	1.70	(1.22)	1.45	(0.92)	0.99	(-0.02)
fac of engineering	1.53	(1.01)	1.31	(0.69)	1.03	(0.09)
fac of agriculture	5.13***	(3.59)	3.26***	(2.87)	2.36***	(2.80)
fac of medicine	3.57***	(3.40)	2.30**	(2.33)	1.68*	(1.91)
fac of pharmacy	3.95***	(3.03)	2.17*	(1.89)	1.58	(1.48)
rank						
rank 1 (t-1)	0.09***	(-10.27)	0.19***	(-7.49)	0.41***	(-4.69)
rank 2 (t-1)	0.19***	(-9.02)	0.40***	(-5.47)	0.60***	(-3.67)
rank 3 (t-1)	0.34***	(-6.88)	0.60***	(-4.04)	0.71***	(-3.52)
other rank (t-1)	0.48***	(-2.96)	0.67**	(-2.34)	0.79	(-1.58)
seniority in rank	0.99	(-0.82)	0.98	(-1.30)	0.99	(-1.41)
head of unit (t-1)	1.26*	(1.93)	1.18	(1.63)	1.13	(1.55)
fulltime at university	4.26***	(3.92)	2.98***	(3.20)	2.68***	(3.20)
tenured	0.69	(-1.40)	0.84	(-0.75)	0.85	(-0.88)
entry as professor >= 1992	1.31	(1.35)	1.51***	(2.68)	1.30**	(2.11)
seniority as professor	1.03	(1.54)	1.04***	(2.61)	1.03***	(2.65)
teaching load	0.93***	(-3.90)	0.98*	(-1.67)	0.98*	(-1.85)
nr previous top performances			1.47***	(15.78)		
nr previous top perf's = 1					13.76***	(22.80)
nr previous top perf's = 2					9.36***	(15.16)
nr previous top perf's = 3					16.33***	(19.46)
nr previous top perf's = 4					16.50***	(17.01)
nr previous top perf's = 5					18.34***	(19.29)
nr previous top perf's = 6					18.14***	(17.60)
nr previous top perf's = 7					21.41***	(21.09)
nr previous top perf's = 8					19.78***	(16.54)
N	5594.00		5594.00		5594.00	
ll	-4784.27		-4601.45		-4393.62	
chi2	440.87		830.61		1645.66	

base categories: rank4 (rank), faculty of physical educ. & kinesiology (faculty membership), agriculture & environment (main discipline), previous top perf's = 0 (event counters)

* p<0.10, ** p<0.05, *** p<0.01

Table 9: Cox models for repeated top performance with frailty

Coefficient	Cox for repeated events, w/ prev. events counter		Cox for repeated events w/ previous events counter and individual frailties		Cox for repeated events, w/ prev. events counter interacted with gender and individual frailties	
	Haz. Ratio	z	Haz. Ratio	z	Haz. Ratio	z
male	1.97**	(2.17)	2.02**	(2.33)	3.11***	(3.27)
age	0.93	(-0.92)	1.01	(0.11)	1.00	(0.03)
age squared	1.00	(0.17)	1.00	(-0.86)	1.00	(-0.80)
main discipline						
biosciences	1.91**	(2.57)	2.56**	(2.40)	2.70**	(2.56)
chemistry	3.00***	(5.08)	6.15***	(5.09)	5.68***	(4.98)
engineering	3.80***	(5.19)	8.74***	(5.23)	8.26***	(5.21)
geosciences	3.26***	(4.24)	9.07***	(4.08)	8.59***	(4.10)
mathematics	4.36***	(5.90)	9.57***	(4.86)	9.19***	(4.89)
medicine I	2.44***	(3.90)	3.51***	(3.37)	3.81***	(3.61)
medicine II	3.07***	(4.95)	4.63***	(4.24)	4.99***	(4.46)
neuroscience	3.47***	(3.48)	9.12***	(3.84)	9.25***	(3.95)
physics	4.10***	(5.13)	8.85***	(5.32)	8.46***	(5.34)
biomedical	2.40**	(2.57)	3.54***	(2.66)	3.73***	(2.82)
biology	1.34	(0.66)	2.30*	(1.87)	1.67	(1.15)
faculty membership						
fac of science	1.45	(0.92)	1.60	(0.97)	1.71	(1.13)
fac of engineering	1.31	(0.69)	1.29	(0.51)	1.39	(0.67)
fac of agriculture	3.26***	(2.87)	5.90***	(3.41)	6.34***	(3.60)
fac of medicine	2.30**	(2.33)	3.27***	(2.90)	3.18***	(2.88)
fac of pharmacy	2.17*	(1.89)	2.83*	(1.94)	3.05**	(2.12)
rank						
rank 1 (t-1)	0.19***	(-7.49)	0.22***	(-5.56)	0.21***	(-5.69)
rank 2 (t-1)	0.40***	(-5.47)	0.41***	(-4.54)	0.39***	(-4.78)
rank 3 (t-1)	0.60***	(-4.04)	0.55***	(-3.66)	0.54***	(-3.73)
other rank (t-1)	0.67**	(-2.34)	0.67*	(-1.82)	0.67*	(-1.89)
seniority in rank	0.98	(-1.30)	0.98	(-0.88)	0.99	(-0.77)
head of unit (t-1)	1.18	(1.63)	1.18	(1.42)	1.16	(1.30)
fulltime at university	2.98***	(3.20)	3.19***	(3.93)	3.13***	(3.91)
tenured	0.84	(-0.75)	0.82	(-0.75)	0.78	(-0.92)
entry as professor >= 1992	1.51***	(2.68)	1.46*	(1.78)	1.37	(1.51)
seniority as professor	1.04***	(2.61)	1.06***	(2.71)	1.05***	(2.58)
teaching load	0.98*	(-1.67)	0.97*	(-1.65)	0.97	(-1.59)
nr previous top performances	1.47***	(15.78)	1.17***	(5.19)	2.51***	(3.99)
male * nr previous top performances					0.47***	(-3.30)
theta			1.58◆		1.42♣	
N	5594.00		5594.00		5594.00	
ll	-4601.45		-4570.36		-4565.33	
chi2	830.61		227.00		246.56	

base categories: rank4 (rank), faculty of physical educ. & kinesiology (faculty membership), agriculture & environment (main discipline)

* p<0.10, ** p<0.05, *** p<0.01

◆Likelihood-ratio test of theta=0: chibar2(01) = 62.18 Prob>=chibar2 = 0.000

♣Likelihood-ratio test of theta=0: chibar2(01) = 53.46 Prob>=chibar2 = 0.000

Table 10: Logit regression results

Coefficient	logistic: high productivity category or not		logistic: persistent top or not (across years)		logistic: persistent good or not (across years)	
	Odds Ratio	z	Odds Ratio	z	Odds Ratio	z
male	2.85***	(2.83)	1.50	(0.62)	2.01*	(1.70)
age	0.96	(-0.34)	1.03	(0.11)	1.19	(0.98)
age squared	1.00	(-0.36)	1.00	(-0.48)	1.00	(-1.42)
main discipline						
biosciences	1.34	(0.64)	1.64	(0.57)	3.17**	(1.99)
chemistry	5.44***	(4.10)	3.84*	(1.74)	9.14***	(3.94)
engineering	8.14***	(4.63)	5.70*	(1.86)	7.01***	(3.02)
geosciences	11.62***	(3.94)	5.64	(1.59)	8.27**	(2.55)
maths	12.25***	(5.00)	2.24	(0.62)	2.50	(0.97)
medicine I	2.62**	(2.49)	2.21	(1.01)	5.75***	(3.23)
medicine II	3.52***	(3.28)	2.74	(1.33)	4.64***	(2.88)
neuroscience	9.26***	(3.62)	4.46	(1.08)	13.55***	(3.00)
physics	7.26***	(4.24)	5.72*	(1.85)	9.34***	(3.42)
biomedical	2.70	(1.58)			6.30**	(2.54)
biology	1.44	(0.46)	0.97	(-0.02)	2.03	(1.02)
faculty membership						
fac of science	1.91	(1.25)	1.88	(0.50)	1.10	(0.12)
fac of engineering	1.65	(0.99)	0.85	(-0.12)	1.16	(0.20)
fac of agriculture	7.83***	(3.44)	4.91	(1.27)	6.30**	(2.40)
fac of medicine	4.87***	(3.60)	2.81	(0.94)	2.90*	(1.74)
fac of pharmacy	7.25***	(3.22)	2.30	(0.55)	1.73	(0.67)
rank♣						
rank 1 (t-1)	0.04***	(-9.89)	0.10**	(-2.44)	0.09***	(-4.01)
rank 2 (t-1)	0.10***	(-8.67)	0.24	(-1.55)	0.21***	(-2.64)
rank 3 (t-1)	0.21***	(-6.88)	0.11***	(-2.81)	0.32***	(-2.86)
other rank (t-1)	0.32***	(-3.12)	0.13***	(-3.47)	0.40***	(-2.59)
seniority in rank	0.97	(-1.12)	0.97	(-0.44)	0.96	(-1.16)
head of unit (t-1)◆	1.44**	(2.02)	1.67	(1.25)	2.62***	(3.91)
fulltime at university♠	5.71***	(4.14)			8.61***	(2.61)
achieved tenure	0.70	(-1.12)	0.23*	(-1.79)	0.45	(-1.32)
entry as professor >= 1992	1.40	(1.26)	0.94	(-0.10)	0.84	(-0.48)
seniority as professor	1.05*	(1.71)	1.08	(1.17)	1.08*	(1.88)
teaching load	0.89***	(-4.01)	0.91	(-1.44)	0.91**	(-2.32)
_cons	0.13	(-0.75)	0.49	(-0.11)	0.00*	(-1.67)
N	5593.00		788.00		991.00	
ll	-1791.73		-150.69		-331.58	
chi2	263.82		49.24		182.75	

base categories: rank 4 (rank), faculty of physical educ. & kinesiology (faculty membership), agriculture & environment (main discipline)

♣ For the latter two specifications, the rank variables indicate the highest rank achieved.

◆ For the latter two specifications, the variable indicates whether the status was ever achieved.

♠ For the persistent top regression, no part-timers are in the persistent top group.

* p<0.10, ** p<0.05, *** p<0.01

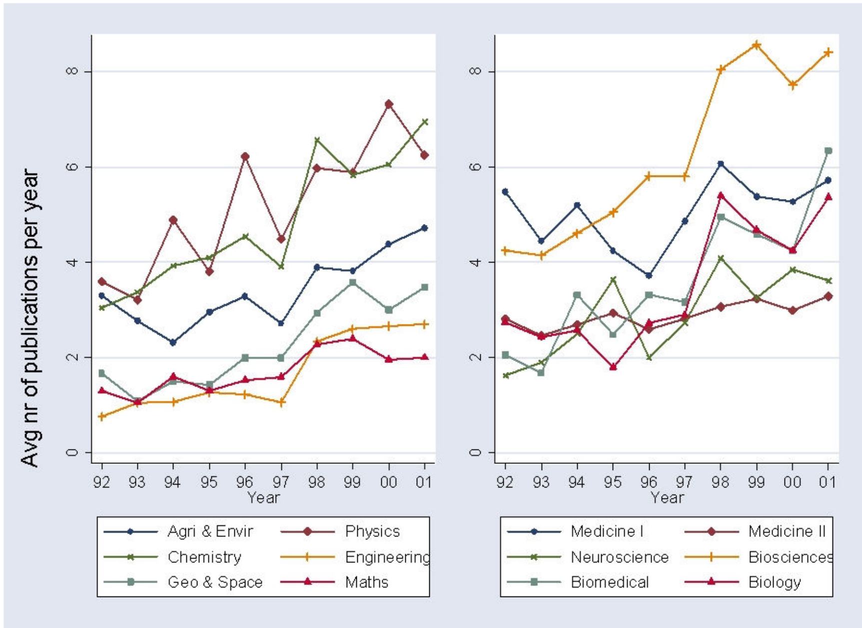

Figure 1: Evolution of publication activity by discipline (1992-2001)

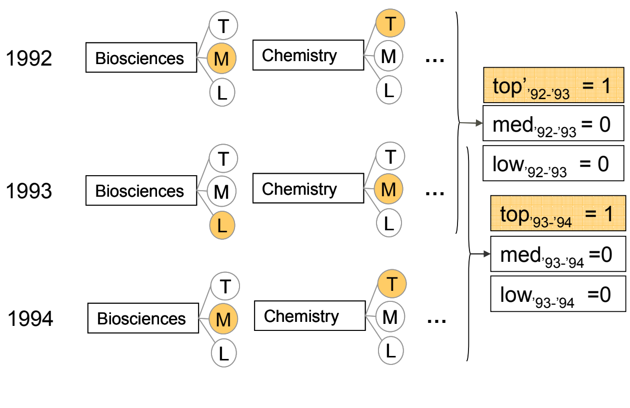

Figure 2: Evaluation of researchers' output

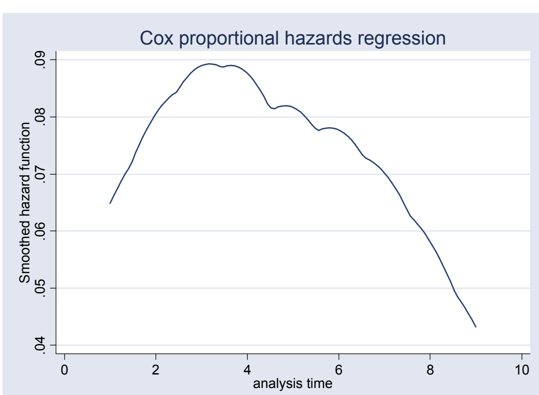

Figure 3: Baseline hazard for repeated top performance

References

Allison, P. D. and J. A. Stewart, "Productivity differences among scientists: evidence for accumulative advantage," *American Sociological Review* 39 (4): 596-606 (1974).

Andersen, P. K. and R. D. Gill, "Cox's regression model for counting processes: a large sample study," *Annals of Statistics* 10: 1100-1120 (1982).

Bernier, C., W. Gill, and R. Hunt, "Measures of Excellence of Engineering and Science Departments: A Chemical Engineering Example," *Chemical Engineering Education*: 194-197 (1975).

Cole, S., "Age and scientific performance," *American Journal of Sociology* 84 (4): 958-977 (1979).

Cole, S. and J. Cole. 1973. *Social stratification in science.* Chicago: University of Chicago Press.

Cox, D. R., "Regression Models and Life Tables," *Journal of the Royal Statistical Society* B34: 187-220 (1972).

Cox, D. R. and D. Oakes. 1994. *Analysis of survival data.* London: Chapman and Hall.

Dasgupta, P. and P. A. David, "Toward A New Economics of Science," *Research Policy* 23 (5): 487-521 (1994).

Jovanovic, B., "Selection and the Evolution of Industry," *Econometrica* 50 (3): 649-670 (1982).

Levin, S. G. S. and P. Stephan, "Research productivity over the life cycle: Evidence for academic scientists," *American Economic Review* 81 (1): 114 (1991).

Long, J. S., P. D. Allison, and R. McGinnis, "Rank Advancement in Academic Careers: Sex Differences and the Effects of Productivity," *American Sociological Review* 58 (5): 703-722 (1993).

Mairesse J. and L. Turner, "Individual Productivity Differences in Public Research : How important are non-individual determinants? An Econometric Study of French Physicists' publications and citations (1986-1997)," *Cahiers de la MSE.Série Verte.* V02066 (2002).

Merton, R. K., "The Matthew Effect in Science. The reward and communication systems of science are considered.," *Science* 159: 56-63 (1968).

Merton, R. K. and H. A. Zuckerman, "Patterns of Evaluation in Science: Institutionalization, Structure and Functions of the Referee System," *Minerva* 9 (1): 66-100 (1971).

Pakes, A. and R. Ericson, "Empirical Implications of Alternative Models of Firm Dynamics," *Journal of Economic Theory* 79 (1): 1-46 (1998).

Prentice, R. L., B. J. Williams, and A. V. Peterson, "On the regression analysis of multivariate failure time data," *Biometrika* 68 (2): 373-379 (1981).

Stephan, P. E., "The economics of science," *Journal of Economic Literature* 34 (3): 1199-1235 (1996).

Stephan, P. E. and S. G. Levin. 1992. *Striking the mother lode in science : the importance of age, place, and time.* New York: Oxford University Press.

Turner, L. and Mairesse J., "La recherche publique dans la production de connaissances, contributions en économie de la science", Université Paris 1 Sorbonne, 2003).

Turner, L. and J. Mairesse, "Mesure de l'intensité de co-publication dans les réseaux scientifiques et évaluation du rôle de la distance géographique," *Revue d'Economie Politique* 2 (2002).

Appendix 1: Constructing the database

The publication and citation data originate from the Science Citation Index (SCI) of the Institute of Scientific Information (ISI). As there is no one-to-one matching between authors and their affiliation address in the ISI data, publications in each of the ten yearly publication files were initially retained when at least one author was affiliated with the KU Leuven so that a number of non-KU Leuven related authors remained present. In a second step, we narrowed the number of publication records by means of a merge with the KU Leuven personnel files in order to only retain KU Leuven affiliated authors (see infra).

Since the ISI records do not allow distinguishing between primary- and co-authorship, we used a `full integer' counting scheme for calculating the performance data. This means that a publication was counted as "1" for all authors of the paper. The same goes for citations: the full number of citations was added to the credits of each author of the paper, whereby an author was identified by his or her surname plus the first initial. This gives rise to homonyms: within the yearly publication files it is not possible to distinguish between authors that have the same name. As discussed below, most of these homonyms could be resolved during the merge with the university personnel file.

Because a researcher's last name plus the first initial is the only piece of information that is shared between the ISI publication records and the university personnel file, the two datasets were combined using this key. In this way, the non-KU Leuven affiliated researchers that were still present in the dataset but whose name did not occur in the personnel files were filtered out. Before carrying out this merge operation, 45 homonyms were dropped from the KU Leuven personnel file since we could not distinguish between these staff members. However, this does not completely rule out mistakes due to homonyms during the merge of the two files. In particular, although homonyms were removed from the personnel file, it is still possible that a homonym occurs between a KU Leuven affiliated author and an external authors within the publications file. Because the name occurs in the personnel file, the publication data of both these author records will be mapped on the staff record, biasing upwards the performance of the staff member. Because the ISI records do not allow linking authors unambiguously to their affiliation, this problem cannot be resolved nor can its magnitude be estimated. We deem it to be a minor issue though, and point out that the merge key used inherently mitigates the problem since researchers with identically spelled last names but a different initial do not yield a 'false positive'.